SIP
戦略的イノベーション創造プログラム

「The second phase of Cross-ministerial Strategic Innovation Promotion Program / Automated Driving for Universal Services / Research on the recognition technology required for automated driving technology (levels 3 and 4) 」

(AD-URBAN Project*)

FY 2018- 2021 Report

Kanazawa University
Chubu university
Meijo university

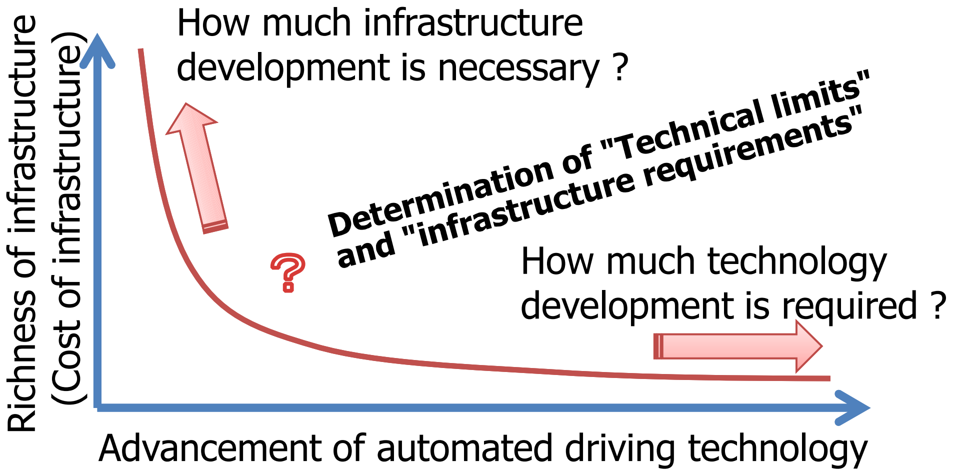
February 2023

---

*FOT project of Automated Driving system under Real city environment based on Academic Researcher’s Neutral knowledge

Overview of this research

- Level 4 equivalent automated driving at urban area
    - It is necessary to have advanced perception and decision-making system by onboard AI, as well as infrastructure such as road facilities and communication facilities to support it
- State-of-the-art automated vehicle technology
    - Knowledge of academia is essential
        - Active collaboration with other projects in SIP



Kanazawa, Chubu, Meijo university

Open research system of university

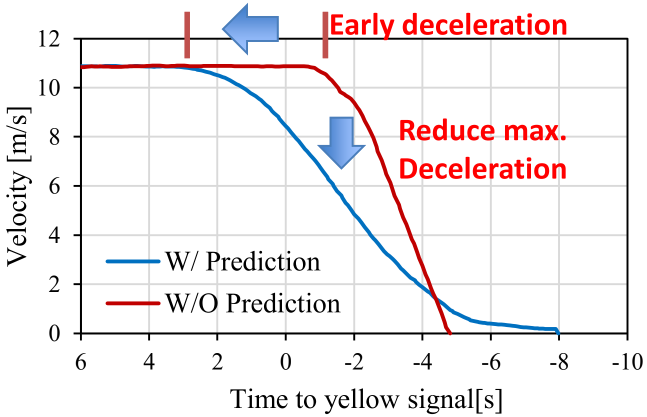
Public road experiment in Tokyo waterfront area

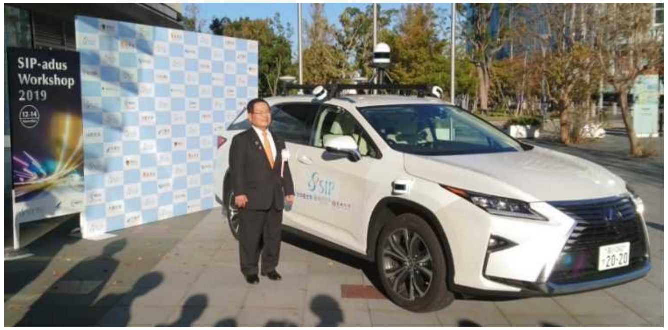
Determination of technical and infrastructure requirements

Overview of Field Operational Tests

- Purpose
  - Verify limitation of recognition technology using onboard sensors
  - Evaluate effectiveness of V2I/V2N
    - Traffic light state and remaining time via V2I/V2N
    - Emergency vehicle position via V2N
  - Evaluate effectiveness of DIVP® simulator
    - Consistency verification for recognition systems
    - Evaluation of performance limitation of recognition systems
  - Provide experimental driving data
    - AD-URBAN Open image dataset v1
      https://github.com/AdmoreKanazawa/open_data

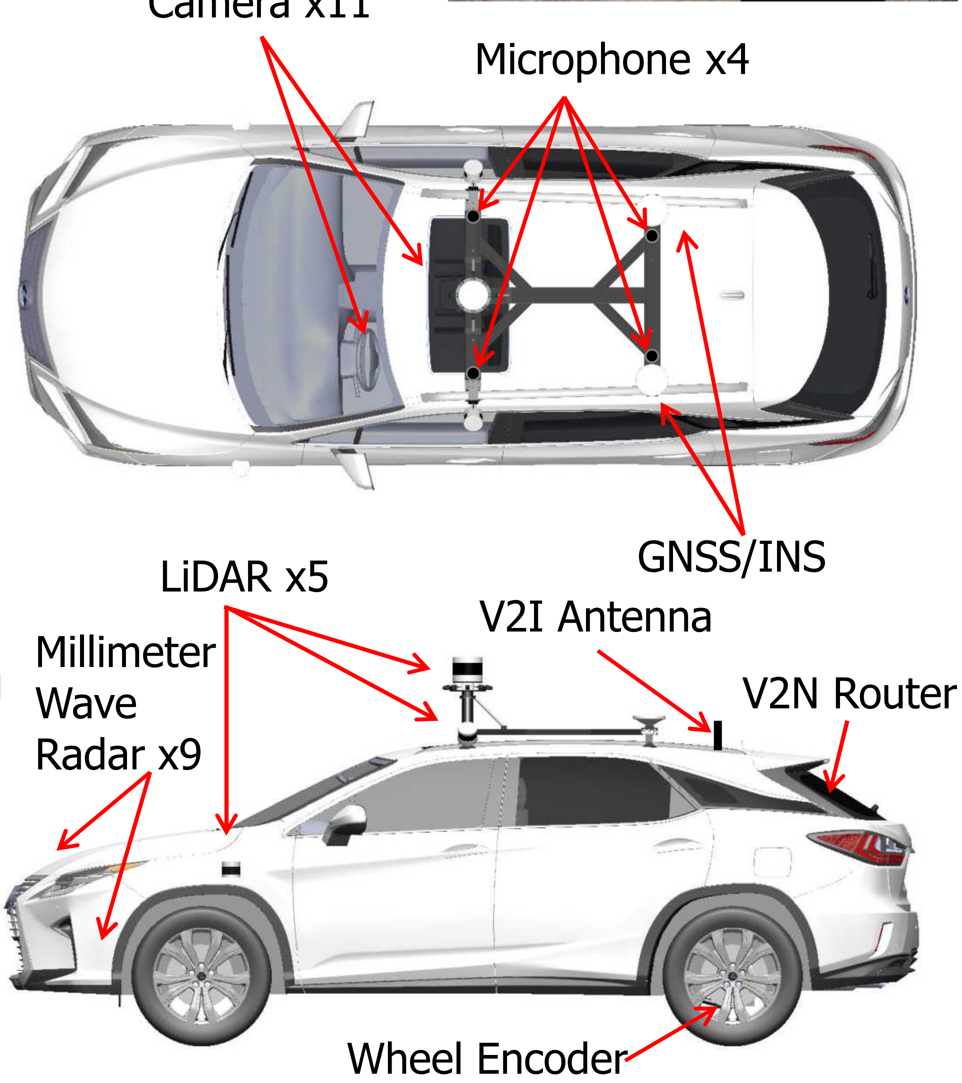
- Activity for improving social acceptability
  - Vehicle exhibition, providing test ride opportunities





Field Operational Tests

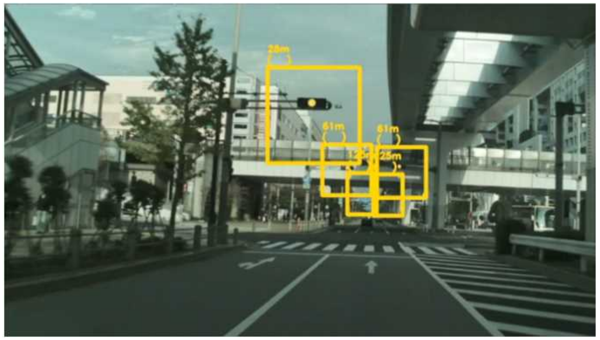
- Development of two test vehicles
    - Tokyo waterfront area
    - Center area of Kanazawa City
- Public road testing in central Kanazawa city
    - Jul. 2019 ~ Feb. 2023
    - Autonomous vehicle
- Public road testing in Tokyo waterfront area (Odaiba and Haneda area)
    - Sep. 2019 ~ Dec. 2022
    - V2I/V2N assisted automated driving + Autonomous vehicle
- Driving record at Tokyo waterfront area
    - 244 days of public road testing
    - Totally 3970.4km of automated driving

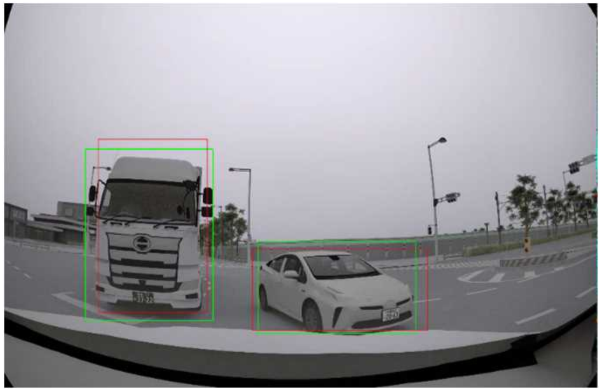


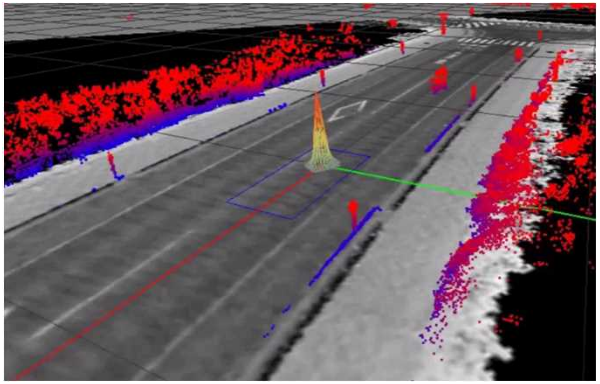
R&D Items and major achievements

Discuss software and hardware infrastructures to improve robust recognition technologies in automated driving

- Traffic light recognition technology
  - Traffic light detection using onboard camera
    - Discussion on infrastructure contributing to improve traffic light recognition rate
    - Effectiveness evaluation using V2I/V2N
- Emergency vehicle recognition technology
  - Siren sound and sound direction recognition using onboard microphone
    - Effectiveness evaluation using V2N
- Object detection technology
  - Object detection using Camera, LiDAR
    - Verification of improvements utilizing digital map for far object detection
    - Effectiveness evaluation using DIVP® simulator
- Self-localization technology
  - Utilization of QZSS “MICHIBIKI”
  - Map matching technology utilizing road paint
    - Discussion on infrastructure contributing to improve self-localization accuracy

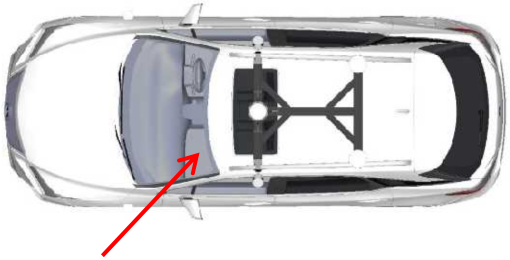






Sensors and technologies for Traffic light recognition

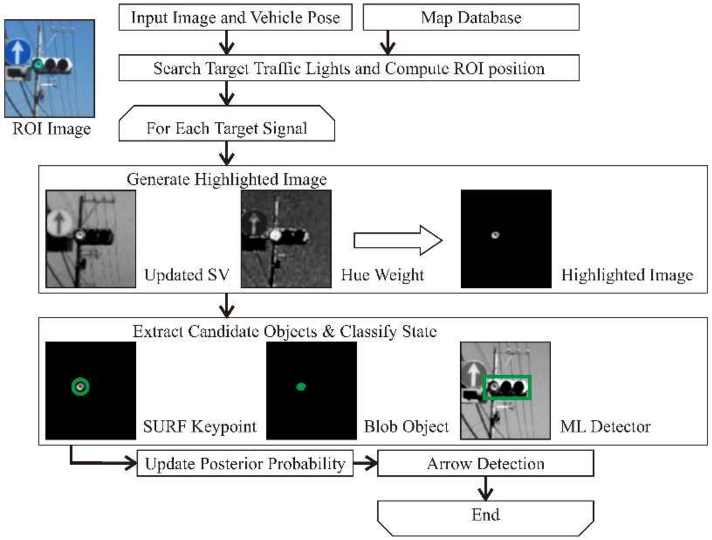
- Sensors
  - SONY IMX390（onboard camera）
    - Full HD resolution
    - LFM(LED Flicker Mitigation) and HDR (High Dynamic Range) functions
- Overview of technologies
  - Lighting object detection with limited recognition area
    - Utilizing traffic light position from digital map with reference to accurate vehicle position based on self-localization
  - Blurred Arrow object detection
    - Machine learning method
    - Consider relative position between traffic light and arrow light



Onboard camera x2

･normal camera（FOV53deg）

･telescope camera（FOV27deg）



K. Yoneda, et al., “Robust Traffic Light and Arrow Detection Using Digital Map with Spatial Prior Information for Automated Driving”, Sensors, 2020

Evaluated results and Failure cases for Traffic light recognition

- Evaluated results
  - Driving data: Tokyo waterfront area
    - 42,603 images (LED Traffic light)
      - Traffic light: 81,273, Arrow light: 8,555
      - Occlusion: 8,157
  - Results:
    - F-value：99.0%
(green, red, arrow lights within 120m)

Recognition rate (within 120m)	Each traffic light	Each intersection
green signal	0.984	**0.997**
red signal	0.977	**0.992**
arrow signal	0.908	0.982
average	0.956	**0.990**

- Environmental failure cases for traffic light recognition
  - Occlusion, background assimilation, night, sunshine (confirmed in FOTs (Field Operational Tests))
  - Heavy rainy（confirmed using DIVP® simulator）


occlusion


background assimilation

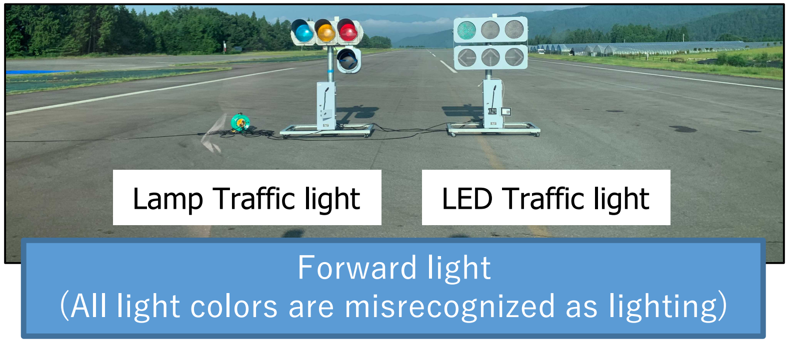

night


backlighting

Discussion on infrastructure contributing to improve traffic light recognition rate

- Considerations based on FOTs results in the Tokyo waterfront area (for LED traffic lights)
  - If recognition failure is temporary, there is no effect on decision to enter intersection
    - Undetected: If multiple traffic lights exist at the intersection, judgment can be made from other visible traffic lights
    - False recognition: If it is a momentary false positive, the effect can be reduced by time series processing
  - If all traffic lights are not detected for a certain period of time,
    - Scenes of undetected traffic lights in the distance are only confirmed, but it is recognizable when approaching an intersection
- Considerations for areas outside the Tokyo waterfront area (for Lamp traffic lights)
  - There are cases where it is impossible to decide to enter intersection due to the influence of forward light
  - There are cases where the recognition distance is shortened when the lighting area of a traffic light appears weak
- Infrastructure conditions contributing to improved traffic light recognition rate obtained from the FOTs
  - It is desirable to have multiple traffic lights at an intersection to avoid failure recognition cases
  - It is more desirable to replace lamp-type traffic lights with LED-type under certain conditions, considering the effects of forward light and the direction of illumination, etc.
  - It is desirable to have traffic lights with wireless infrastructure installed to ensure robustness

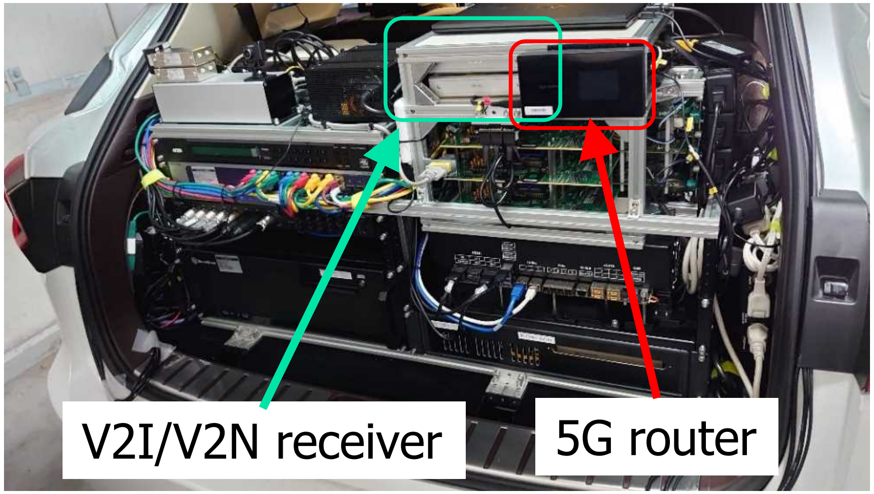
Backlighting
(It can be judged from other visible traffic light)



Forward light
(All light colors are misrecognized as lighting)

Effectiveness evaluation using V2I/V2N

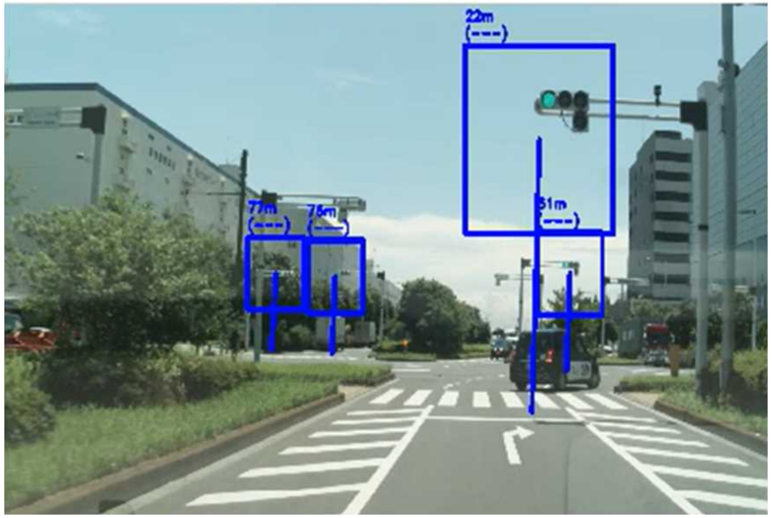
- Hardware used at Tokyo waterfront area
  - Onboard receiver for V2I and V2N
    - Items on loan from the Tokyo Waterfront Area FOT
- Evaluation
  - Compare recognition distance and timing for onboard camera, V2I, and V2N
    - Install both V2I and V2N receiver in automated vehicle and collect driving data and both received information simultaneously
  - Verify effectiveness of reducing deceleration in dilemma zone



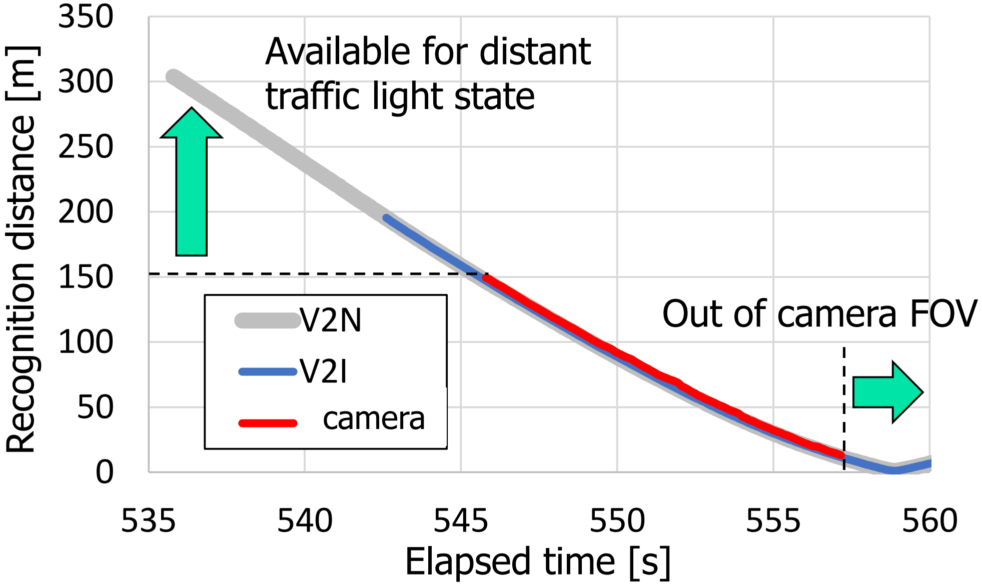
Comparison of traffic light recognition distance and timing for each method

(*) Calculated as distance required for 0.1G deceleration at metered vehicle velocity of 60 km/h (actual vehicle velocity approximately 55 km/h)

- Traffic light recognition distance
  - V2N>V2I>Camera (maximum recognition distance by camera is about 150m)
    - Approximately 120m(*) is required for smooth decision to approach an intersection
    - Recognition distance is acceptable with passing through intersections for all methods
  - Further robustness can be implemented by fusion of all methods
- Traffic light recognition timing
  - There is no significant discrepancies by the human eye
    - Notification timing of recognition results by camera and switching the traffic light state by wireless infrastructure (V2N/V2I)
    - Confirmed in FOTs in Tokyo waterfront area and Nara Prefecture



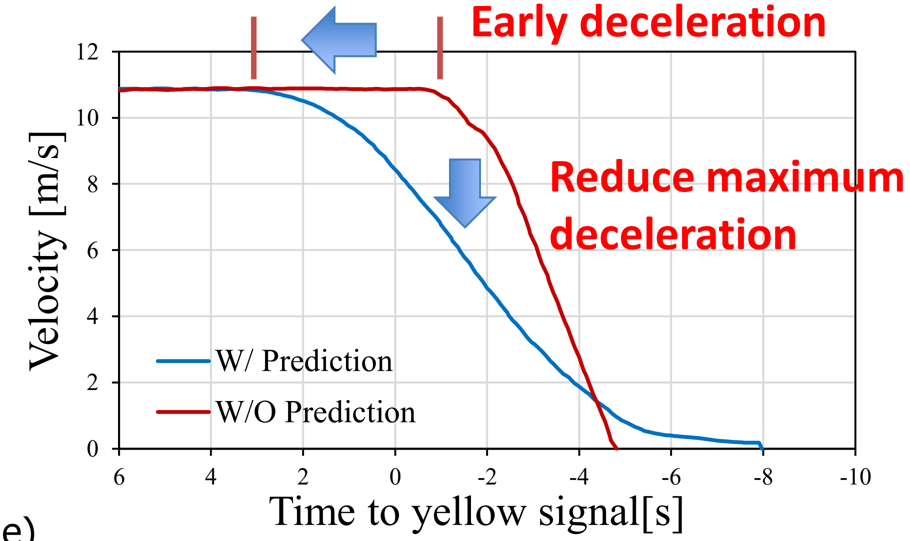
Recognition status when passing the intersection



Effectiveness verification of traffic light-coordinated driving using information on the remaining time via V2I/V2N

- Effectiveness of infrastructure-assisted traffic lights
  - Data acquisition of lighting state information by V2I/V2N
  - Utilizing remaining time information
- Advance deceleration in dilemma zone using the remaining time at traffic lights
  - Public road test: Haneda and Odaiba areas
  - Verify effectiveness of reducing deceleration in dilemma zones on public roads
    - Autonomous intersection entry: -0.4G ⇒ Using V2I/V2N information：-0.2G

Example of V2I Information (Visualize Remaining Time)



Summary of issues in Camera based recognition and effectiveness of V2I/V2N

- Issues in camera based traffic light recognition
  - Achievement of 99% or higher recognition rate (Tokyo waterfront area)
    - red, green, arrow lights within 120m
  - Failure scenes
    - Occlusion, background assimilation, backlighting, night, heavy rainy, etc.
- Effectiveness of traffic light information via V2I/V2N
  - Acquiring of traffic light states out of sight
    - E.g. Occlusion by large vehicle, and Traffic lights after passing through a curve
      - Onboard cameras can decide to entry intersection after seeing the traffic lights as well as humans do
  - Improving robustness by multiple system configuration
    - Especially effective at an intersection with only one traffic light
  - Reducing maximum deceleration in dilemma zones using the remaining time

Occlusion by large vehicle

Traffic lights after passing through a curve

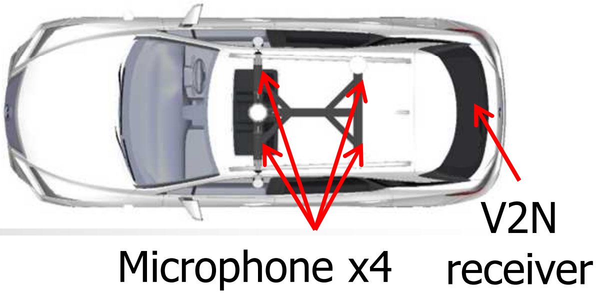
occlusion

background assimilation

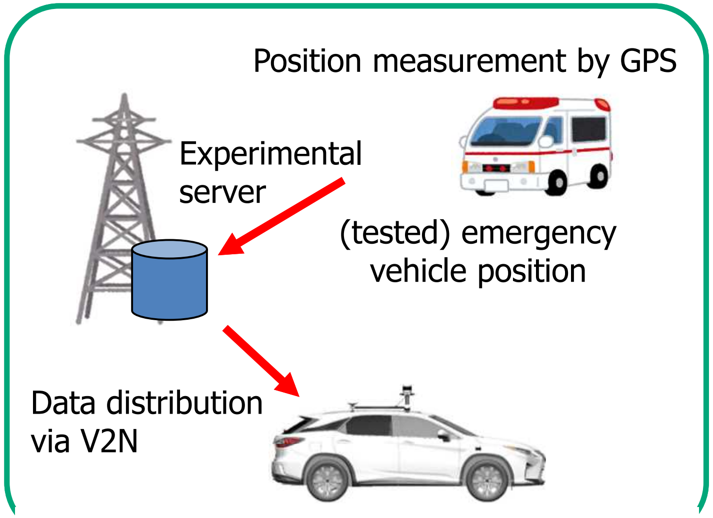
night

backlighting

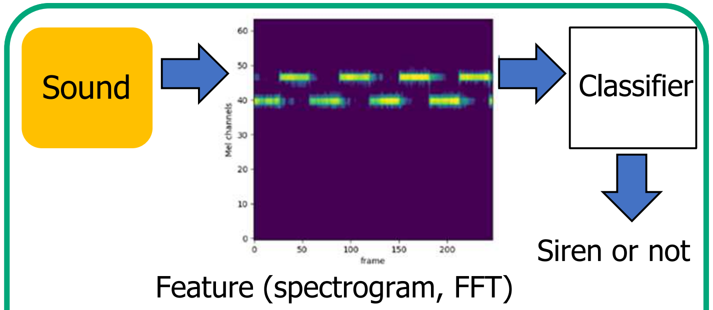
Sensors and technologies for the emergency vehicle recognition



- Sensors
  - V2N onboard equipment
    - Items on loan from the Tokyo Waterfront Area FOT
  - UETAX um-100(waterproof microphone)x4
- Overview of Technologies
  - Tested Emergency vehicle position information via V2N
    - Tested Emergency vehicle position information is measured by GPS
    - Measured position information is distributed via cellular communication
  - Recognition of siren sound and sound direction using microphones
    - Siren recognition by machine learning
      - Sound feature extraction
        - Spectrogram, FFT
      - SVM: Support Vector Machine
    - Sound direction recognition by phase difference between multiple microphones



Emergency vehicle recognition via V2N



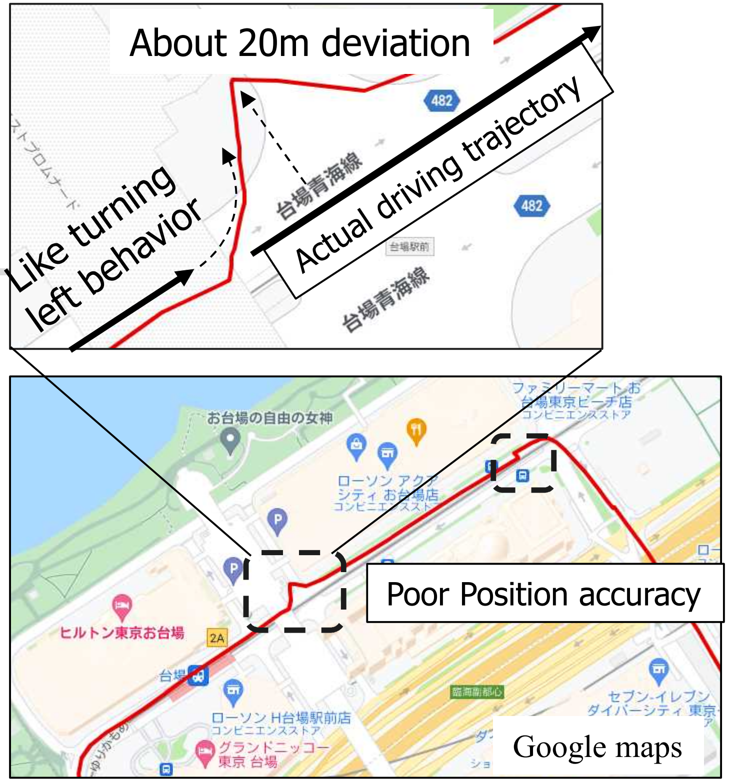
Feature (spectrogram, FFT)

Sound recognition: [C. Takatsu, et al., 2022 (Japanese)] + sound direction recognition

Siren recognition using microphone

Evaluated results of emergency vehicle information and failure cases

- Emergency vehicle position via V2N
  - Evaluation for test vehicle
    - Join FOTs of position data distribution for tested emergency vehicles at the Tokyo waterfront area
  - Evaluated results
    - Smooth driving trajectory can be obtained generally
    - About 20m positioning deviation is occurred depending on the situation.
- Siren recognition using microphones
  - Evaluated data
    - Driving data in Kanazawa city, and on a test course
  - Evaluated results
    - Siren recognition rate: almost 96%
    - Siren direction accuracy: almost 10 deg (in good visibility conditions)
  - Failure cases for recognitions
    - Noise from rain, wind noise, etc.
      - Not fatal issues, but needs to consider installing conditions
    - Reflected sound, and occlusion by buildings, etc.
      - Influence on direction estimation

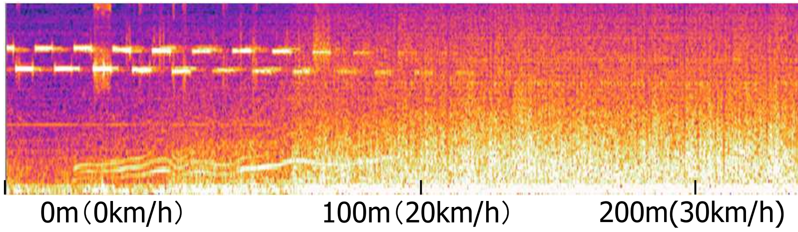


Distributed V2N information 2022/1/13 13:30

Evaluated results for siren recognition

Precision	Recall
0.966	0.960

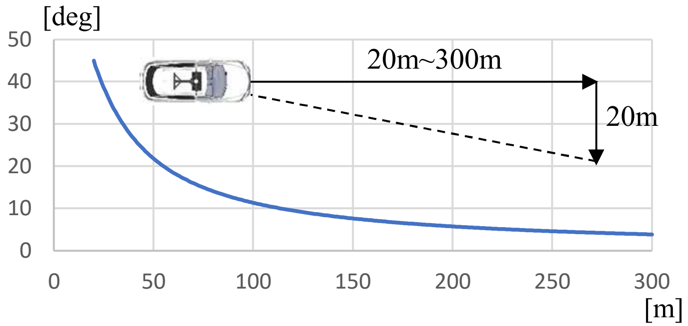
Precision: Index of low false detections
Recall: Index of low miss detections



Spectrogram when an emergency vehicle is stopped and the vehicle accelerates and moves away from a short distance to a long distance

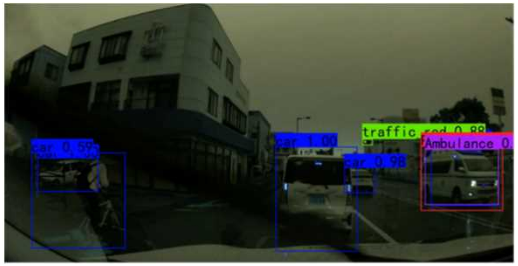
Effectiveness of Emergency Vehicle Position Information via V2N

- Considerations based on the FOTs in the Tokyo Waterfront Area
  - Emergency vehicle location information via V2N
    - Particularly effective in occluded or distant scenes
    - Azimuth accuracy is acceptable at long distances
    - Positioning accuracy is depended on satellite conditions especially at short range
      - Issue: understanding detailed vehicle behavior
  - Siren sound recognition using microphone
    - Siren direction estimation accuracy is about 10 deg (in good visibility conditions)



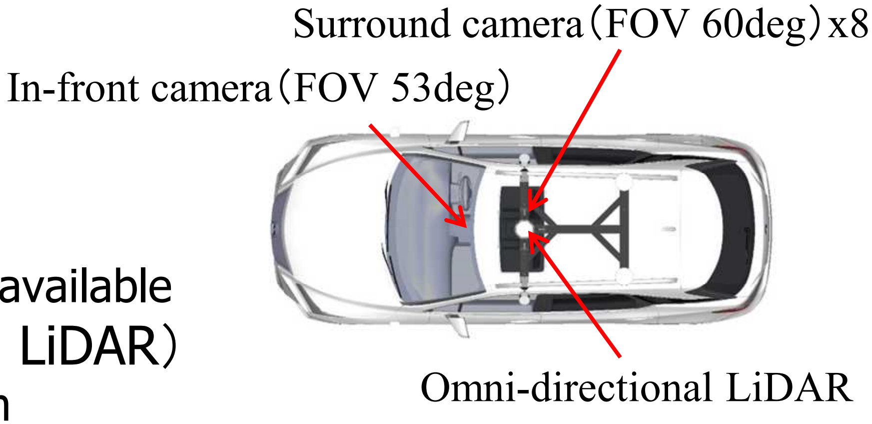
Influence of emergency vehicle position error on azimuth

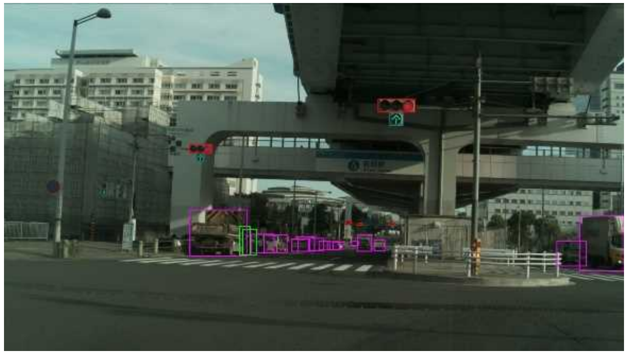
- Effectiveness of emergency vehicle location information via V2N
  - For manual driver or automated driving systems below Level 3
    - If the driver eventually takes avoiding action against an emergency vehicle in a short range, the V2N information is expected to function effectively as a means of providing information at an early stage
  - For automated driving systems of Level 4 or higher
    - It is necessary to take avoiding action according to the driving position of emergency vehicles and surrounding vehicles
    - It is important to understand the detailed behavior of emergency vehicles at a short range
    - It is also necessary to combine methods such as siren sound recognition and image recognition to understand the detailed behavior

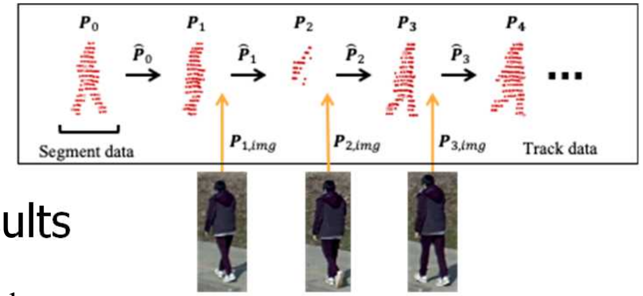


Sensors and Technologies for object detection

- Data Acquisitions
  - SONY IMX390 (in-front camera)
  - Onsemi AR0231 (surround camera)
    - resolution：Full HD
    - HDR（High Dynamic Range） function available
  - Velodyne VLS-128（omni-directional LiDAR）
    - # of lasers：128， max distance：300m
    - FOV: horizontal 360deg, vertical 40deg
- Technologies
  - Object detection (camera)
    - 2D bounding box from DNN*
      - Our trained model based on YOLOv3, YOLOv4
      - Region of Interest (ROI) from digital map
  - Object detection (LiDAR)
    - 3D bounding box from DNN*
      - Our trained model based on PointPillars
  - Fusion both 2D and 3D bounding boxes
    - Improve accuracy by combing LiDAR/Camera results



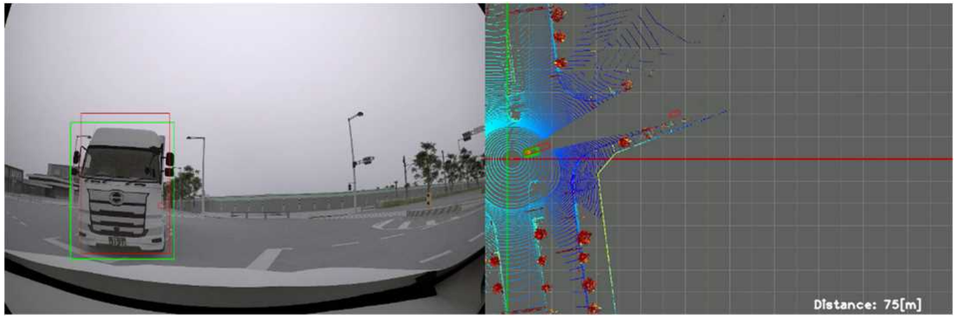




*DNN：Deep Neural Network

Object detection results and failure cases

- Evaluation
  - Data (collected area)
    - Odaiba & Haneda
    - Kanazawa Univ.
  - Results (detected max distance)
    - vehicle： within 200m
    - pedestrian：within 70m
      - Confirm Effectiveness of fusion of LiDAR/camera
  - Failure cases
    - Occlusion, similar background, rainy（camera, LiDAR）
    - Night, backlight（camera）
      - Confirm in FOTs, simulation



Blind spot impact assessment（red：correct, green： result）

Influence of rainy

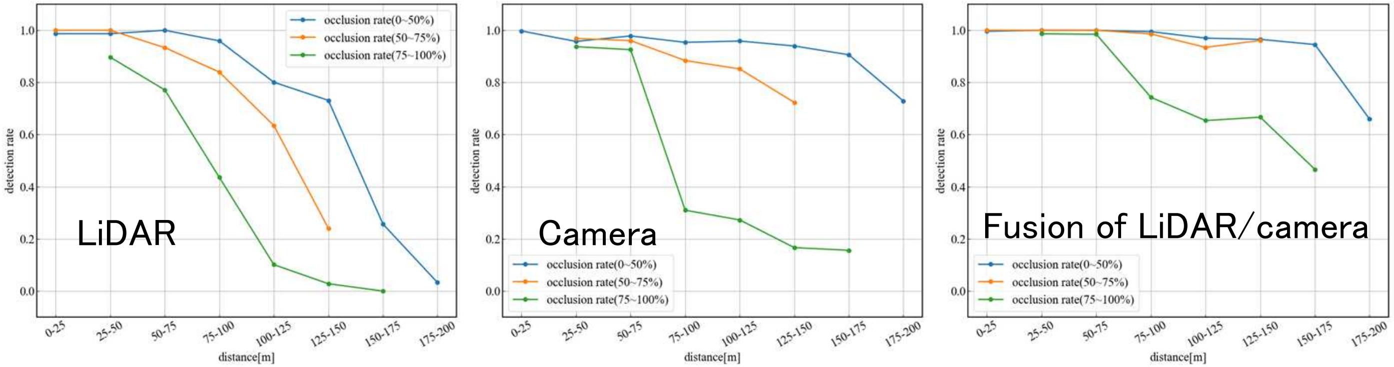
Detection accuracy under different occlusion rate

（This evaluation is done each frame. It expect that the accuracy may improve by using tracking process.）

※definition of occlusion rate:

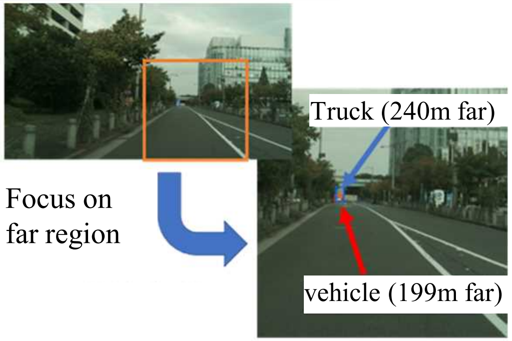
・LiDAR: invisible area ratio of 3D box

・Camera：inviable area ratio of 2D box

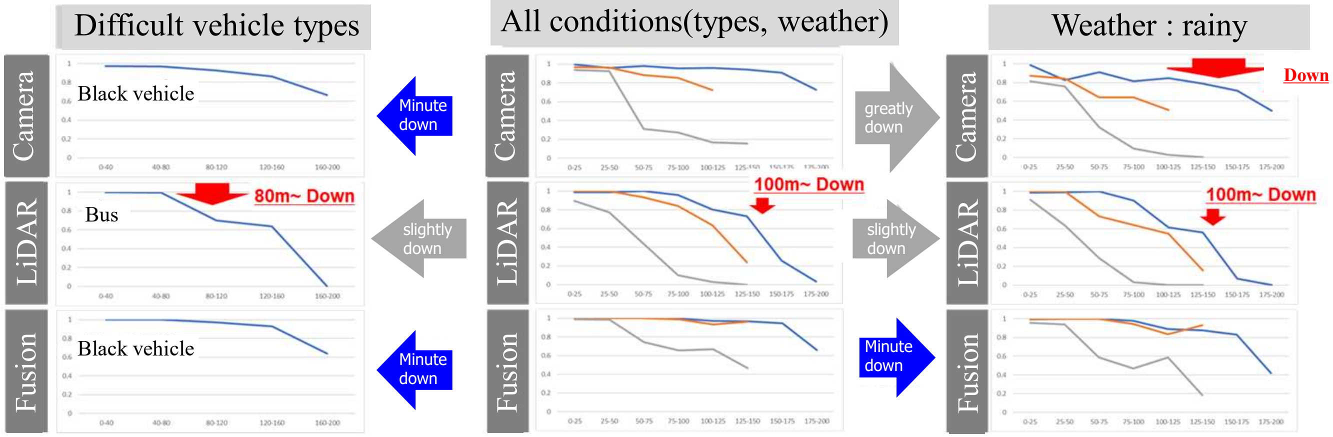


Effectiveness of evaluation using simulation environment in object detection

- Consideration based on the FOTs
  - Distance of detection become shorter by factors
    - occlusion, similar background, rainy, etc.
  - Solution : utilize high precision digital maps
    - To focus on far region, detection process is done only ROI
    - Effective in urban including right-turn in intersection etc.
  - Necessary to clarify the limitation of sensors
    - It needs huge cost and inefficient in real environment
- Effectiveness of evaluation using simulation environment
  - Utilization of simulation is important aspect of reproduce difficult scenes
    - Expect sensors layout and model training to improve the accuracy
  - Simulation environment is necessary for safety assessment



Processed within ROI



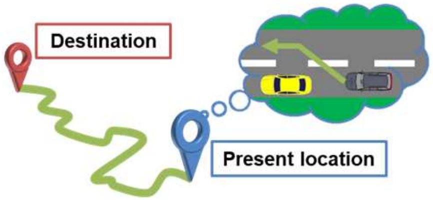
Accuracy of various difficult scenes

(This evaluation is done each frame, it expect that the accuracy may improve by using tracking process.)

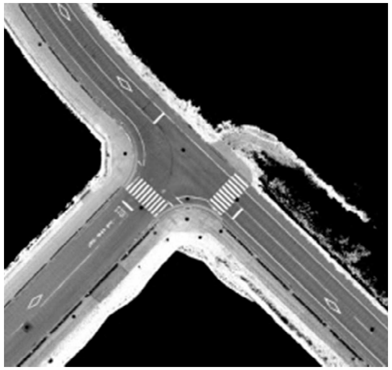
The results reported by collaboration with DIVP®

Self-localization Technology Efforts

- The necessity of self-positioning estimation technology
  - High-precision maps are effective for automatic driving in urban areas
  - High-precision self-positioning is essential to use high-precision maps
    - Difficult to estimate self-position by GNSS alone, e.g., in tunnels
    - Accurate self-position estimation by map-matching
  - Toward Improving Reliability
    - It is important to improve both GNSS/INS* and map-matching



- Specifics of Research and Development
  - 1.Development of GNSS/INS
    - Robust lane-level position estimation (1.5m accuracy)
      - Assumed to be used in conjunction with map-matching technology, etc.
    - RTK-GNSS (0.3m accuracy) reliability estimation
      - Utilization of Quasi-Zenith Satellite MICHIBIKI
      - Level of automatic operation possible with GNSS/INS alone
  - 2.Development of map-matching technology
    - Map-matching technology utilizing ortho-images
    - High-precision self-position and attitude estimation using vehicle-mounted grade GNSS/INS in combination

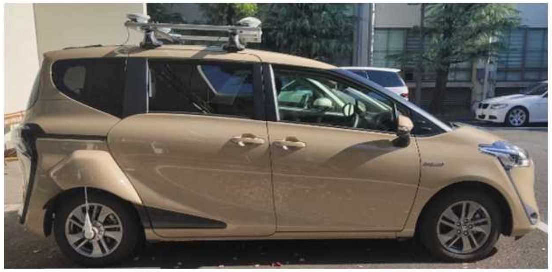


---

*GNSS/INS：Global Navigation Satellite System /Inertial Navigation System

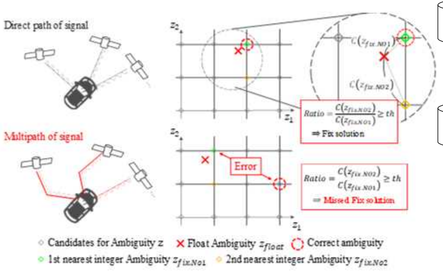
Overview of Satellite Positioning Technology (GNSS/INS)

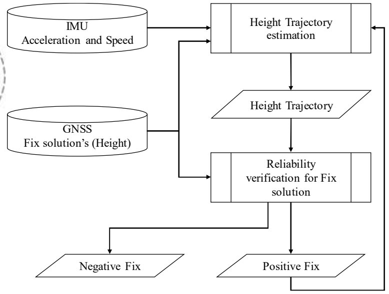
- Sensors
  - Septentrio Mosaic-X5 (MICHIBIKI-compliant GNSS receiver)
    - CLAS-LIB enables PPP-RTK using CLAS
  - ADIS16475 (general-purpose MEMS-IMU)
    - 6-axis MEMS-IMU with a level that can be installed in vehicles in the future
- Overview of technology
  - High accurate position estimation using vehicle motion
    - Accurate positioning by high precision estimation of vehicle motion using GNSS Doppler
  - Improvement of RTK-GNSS usability by utilizing vehicle motion constraint
    - Improvement of RTK-GNSS positioning accuracy by improving the initial position of RTK-GNSS search



GNSS Receiver
mosaic-X5
Septentrio

MEMS-IMU
ADIS16475
AnalogDevises





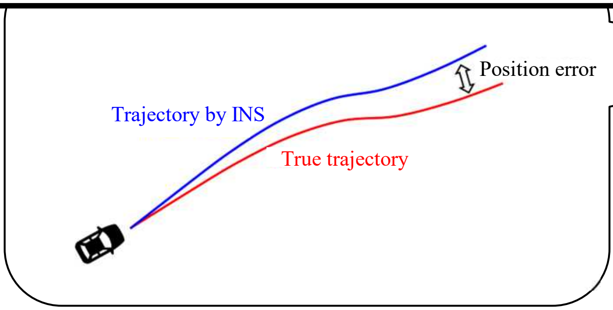
Aoki Takanose, et. al., Improvement of Reliability Determination Performance of Real Time Kinematic Solutions Using Height Trajectory, Sensors, 21, 2, 2021.1

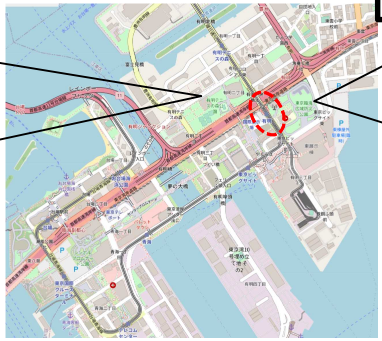
Satellite positioning technology (GNSS/INS) evaluation results and underperformance factors

- Performance evaluation results
  - Evaluation data (Tokyo waterfront area)
    - 10 km route including Yurikamome elevated railway line
  - Evaluation index: Percentage of positional accuracy achieved
    - Achievement rate of 1.5m accuracy, reliability of 0.3m accuracy
    - Percentage of 0.3m accuracy achieved 10 seconds after correction information is lost
  - Evaluation result
    - 1.5m accuracy achieved 96%, 0.3m reliability 99%, 0.3m accuracy achieved 89% after 10 seconds
- Factors causing positional inaccuracy
  - GNSS signal degradation due to buildings and structures
  - Integral position error due to INS error

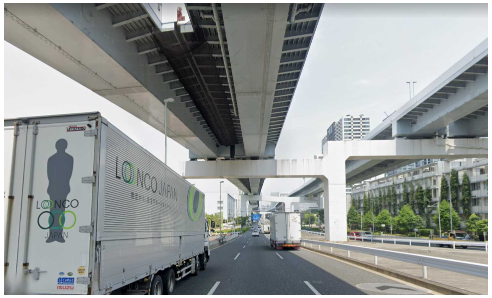
	Percentage
1.5m accuracy	96%
0.3m reliability	99%
0.3m accuracy after 10 seconds	89%

Position error due to INS after interruption of correction information





Under the overpass/GNSS signal degradation

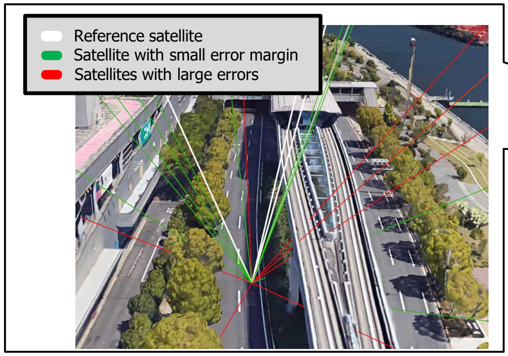


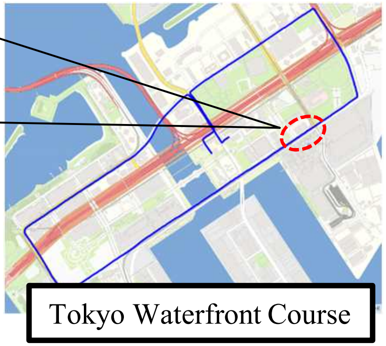
Highly Reliable Self-Positioning Technology Using Satellite Positioning Technology (GNSS/INS)

- Consideration based on the results of verification in the Tokyo waterfront area
  - Consideration of GNSS signal degradation due to buildings and structures
    - Lane-level position estimation (accuracy: 1.5m): generally applicable with improved algorithms (96%)
    - GNSS/INS alone can be used for automated driving (accuracy of 0.3m): Reliability can be determined (99%). However, it is difficult to completely eliminate position estimation errors using GNSS alone
  - Consideration of integral position error due to INS error
    - If absolute position cannot be estimated due to GNSS signal shielding/disruption of correction information, it is difficult to maintain 0.3m accuracy for 10 seconds.
    - Even if a highly accurate INS (FOG*) is used, the position accuracy after 10 seconds is equivalent to that of a general-purpose MEMS-IMU. However, it is highly possible to maintain 0.3m accuracy for a shorter period of time (8 seconds)
- Toward reliable self-position estimation using satellite positioning technology
  - In areas where the number of GNSS satellites decreases or accuracy deteriorates due to buildings/structures, it is particularly important to collaborate with other methods such as map-matching technology.
  - It is also important to identify in advance the sections where positioning accuracy will deteriorate
    - Use of 3D map data, etc

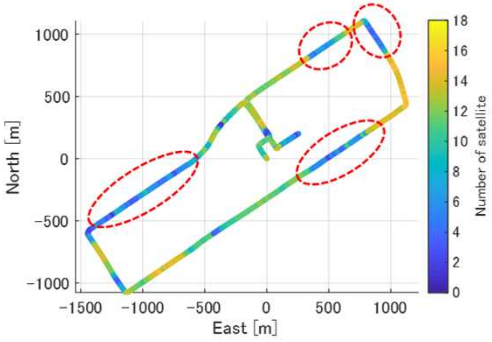
*Fiber Optic Gyro

An example of using 3D map data to gain a preliminary understanding of the environment in which satellite positioning accuracy deteriorates.



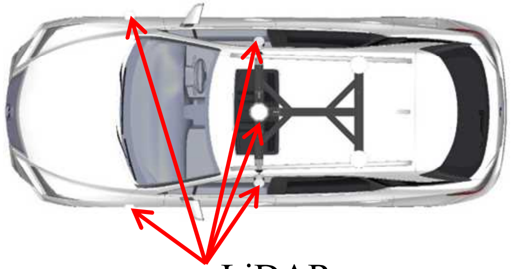


Locations where the number of satellites is reduced

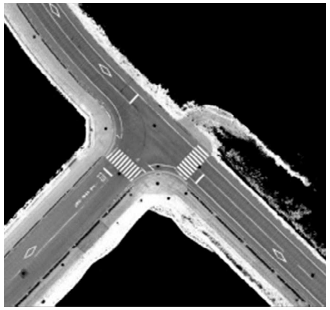


Overview of Map Matching Technologies

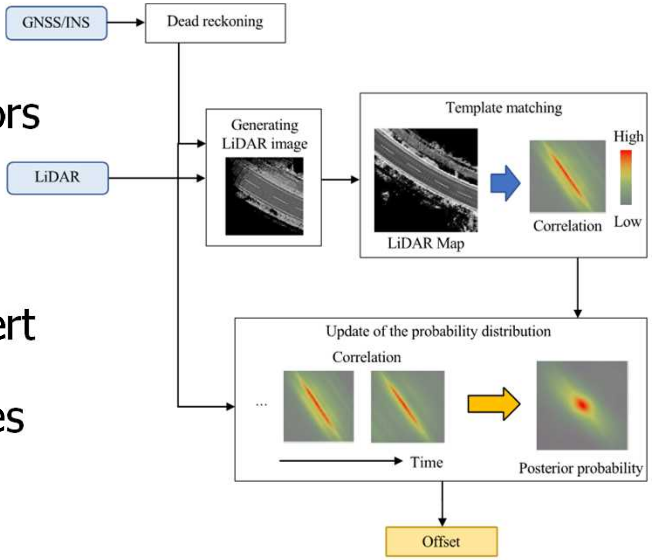
- Sensors
  - GNSS/INS: Applanix POS-LV125
    - GNSS/INS with MEMS-IMU
      - Used for dead reckoning
  - LiDAR: Velodyne LiDAR × 5
    - Used for map matching (Ortho image generation)
- Overview of technology
  - Dead reckoning (DR)
    - Position estimation by integrating velocity vectors
  - Map matching
    - Ortho image generation from LiDAR reflectivity
      - Creating road pattern images
      - Reflectivity correction based on the Lambert model
    - Error correction of DR by matching ortho images
      - Estimating dead reckoning error
    - Improving accuracy and robustness by time series processing



LiDAR
・VLS-128 ×1
・VLP-32c ×2
・VLP-16 ×2



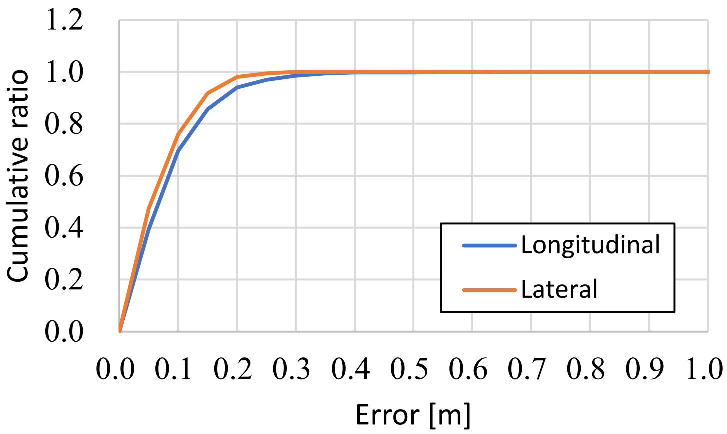
Ortho image
(using LiDAR)



Reflectivity correction model, yaw angle estimation technology, etc. added based on the following research:
N. Suganuma, D. Yamamoto and K. Yoneda, "Localization for autonomous vehicle on urban roads", Journal of Advanced Control, Automation and Robotics, 2015.

Evaluation results and failure cases

- Evaluation of localization accuracy
  - Data for evaluation in Kanazawa city
    - Driving route of approximately 20 km
    - Route that includes urban and mountain areas
  - Evaluation index
    - Position error calculated using the GNSS/INS post-processing results as the ground truth
  - Results
    - Longitudinal RMS error 0.097 m and lateral RMS error 0.079 m



- Failure cases
  - Material of road surface (thermal barrier coating and concrete pavement)
  - Blurring of road patterns and wet road surface cased by rain

Camera image

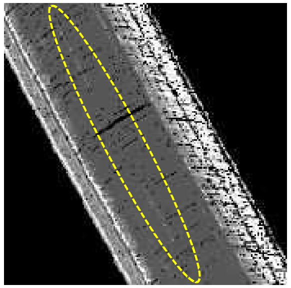
LiDAR image (ortho map)

Road with thermal barrier coating

Camera image

Wet road with concrete pavement

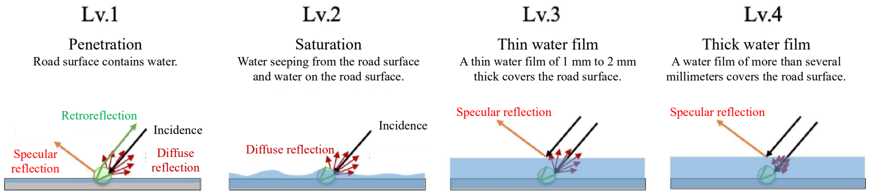
LiDAR image



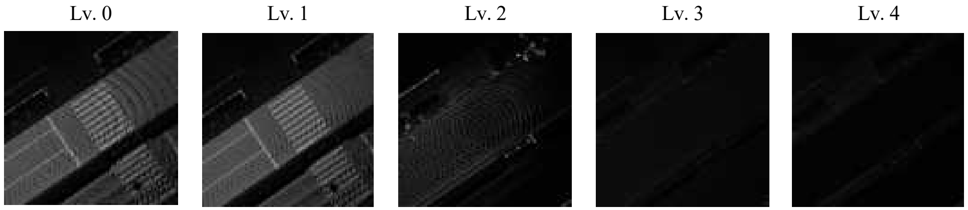
Fading of dashed line

Conditions of infrastructure for improving localization accuracy by map matching

- Consideration based on the FOTs (using map matching with ortho images)
  - Highly accurate localization can be achieved using map matching technology, including in places where GNSS/INS positioning accuracy is reduced (e.g., in tunnels and under overpasses).
  - Road surface wetness affects localization accuracy in areas where white line contrast is low.
    - Thermal barrier painted roads and concrete paved roads (Evaluation of performance limits using virtual space has also been conducted)
  - Sections with unclear road patterns affect the localization accuracy.
    - Blurring of white lines (The effect of blurriness on recognition has also been evaluated)
  - However, if the interval affecting localization accuracy is short, there is no problem in practical use.
    - FOTs areas in Tokyo Waterfront Area and Kanazawa City
- Conditions of infrastructure for improving localization accuracy concluded from the FOTs. It is recommended to avoid situations where the infrastructure for localization (e.g., white lines) doesn't exist for a long continuous distance
  - Infrastructure for localization had better be easy to distinguish from other patterns (e.g., high contrast between white lines and the pavement)
  - Considering the malfunctions of map matching, multiplexing techniques are needed, for example, other map matching techniques or combination with GNSS/INS.



Definition of road surface wetting level



Effect of wet road surface with thermal barrier coating
(Investigation of performance limits by simulation)

Summary

- FOTs at Tokyo waterfront area
  - Period: September 2019 ~ December 2022
  - Driving days: 244 days
  - Automated driving distance: 3970.4 km
- Reports on the hardware and software infrastructures contributing to the improvement of recognition technologies in automated driving
  - Traffic light recognition technology
  - Emergency vehicle recognition technology
  - Object detection technology
  - Self-localization technology
    - Satellite positioning technology utilizing QZSS “MICHIBIKI”
    - Map matching technology using road paints
- Toward future utilization of research and development results of AD-URBAN Project
  - We will continue to share information on the status of infrastructure that contributes to the improvement of automated driving technology with related parties as necessary.
  - We will continue to provide data obtained from FOTs (AD-URBAN Open image dataset v1), even after the completion of the project.
  - We will continue to discuss with DIVP®, SAKURA, and other related parties for efficient and comprehensive safety evaluation of automated driving technology, even after the completion of the project

This report documents the results of Cross-ministerial Strategic Innovation Promotion Program (SIP) 2nd Phase, Automated Driving for Universal Services (SIP-adus, NEDO management number: JPNP18012) that was implemented by the Cabinet Office and was served by the New Energy and Industrial Technology Development Organization (NEDO) as a secretariat.